

Weekly Report

Lu Junhua

2016 年 6 月 20 日

Done

- Finished reviewing the vinci paper.
- Finished one course essays (two remaining), took an exam (one remaining).
- Interviewed candidate students of graduation design.

To do

- Finish two course essays, translate one paper and review & take an exam.
- Discuss about the reviewers' comments on our VIS paper.

论文

- 15infovis *Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations*. 文章的核心是要做到breadth-oriented的可视探索, 而不是depth-oriented的. Data-variation (不同的变量选择和transformation)优先于design-variation (相同数据的不同的可视编码). 本身文中尽管只是对多变量的表格数据进行探索, 但算是挖了一个坑, 也许他们已经有了相关的产品雏形(因为是Jeff Heer 那边的). 其整个系统工作流程如下图 1, 其中的Compass作为推荐引擎, 里面有重要的一步是通过perceptual effectiveness score来帮助选择较好的可视编码, 并修剪掉近似的结果防止视觉混乱. 这一套工作对viscomposer或者类似的g2有一定的启发意义, 文中还有一套语法(不过也是借鉴了别人的)而对于我们一般的工作来说, 可视推荐在选择编码中有提高探索效率的意义, 因为在探索的时候, 如果既要选择编码, 又要在编码好以后进行分析, 相当于额外负担.
- 16pvis *Interactive Visual Co-Cluster Analysis of Bipartite Graphs*. 最近见到了三次关于co-cluster或者叫bi-cluster的可视化. 这篇文章一开始提到了bi-cluster可以看作是二部图两边分别聚类, 这是我之前没有想到一个概念. 我暂时能理解的可能就是co-occur或者协同过滤中可能会较多的用到这个东西. 这篇文章贡献点在于先提出了一个半监督的bicluster算法, 然后可视交互辅助这个算法. 可视编码中利用邻接矩阵和树图的形式来展示cluster内部的元素关系, 如图 2 而用连线展示不同cluster 之间的联系. 我个人对bicluster也不是特别了解, 我估计大概和ego-network差不多, 是图可视化中一个小分支, 只是近来可能有较多的应用, 包括我们后面会做的异常检测, 作弊检测等.
- 15scivis *The Data Context Map: Fusing Data and Attributes into a Unified Display*. 尽管是一篇scivis的文章, 其理念仍然可以借鉴. 文章把维度和数据相似性一同考虑进来. 这个问题我们可能曾经在哪里见过, 这篇文章做出来觉得蛮对的. 文章从一个择校问题引入, 并用同样的例子来解释, 容易理解. 可视化过程中结合了一些改进的mds 和kde进行布局与绘制.

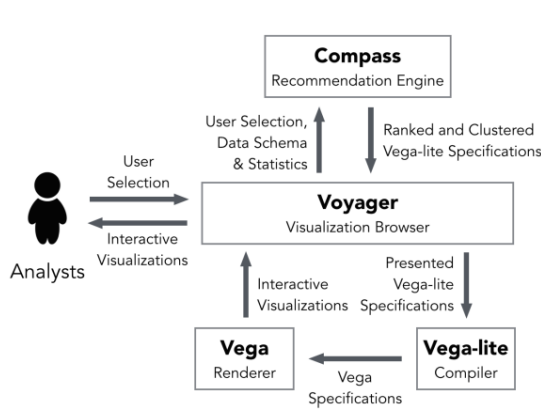


图 1: Voyager系统流程

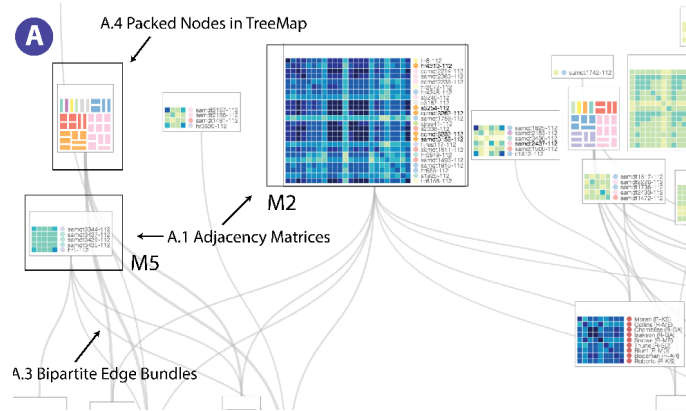


图 2: bi-cluster的可视编码

- 16eurovis *TimeArcs: Visualizing Fluctuations in Dynamic Networks*. 该工作与审的那篇VINCI有一定类似之处(图或者网, 差异). 其基本思想是结合了arc diagram, 流图, 并用力引导来布局. 力引导将相似的放一起, 如图 3, 一定程度上减少了visual clutter, 这一点以前并没有意识到. 交互方面没有太多的先进技术, 这个方案用于展示还是有一定美观性的, 如果用于发现一些潜在的规律也可以, 但可能需要增加一些交互.

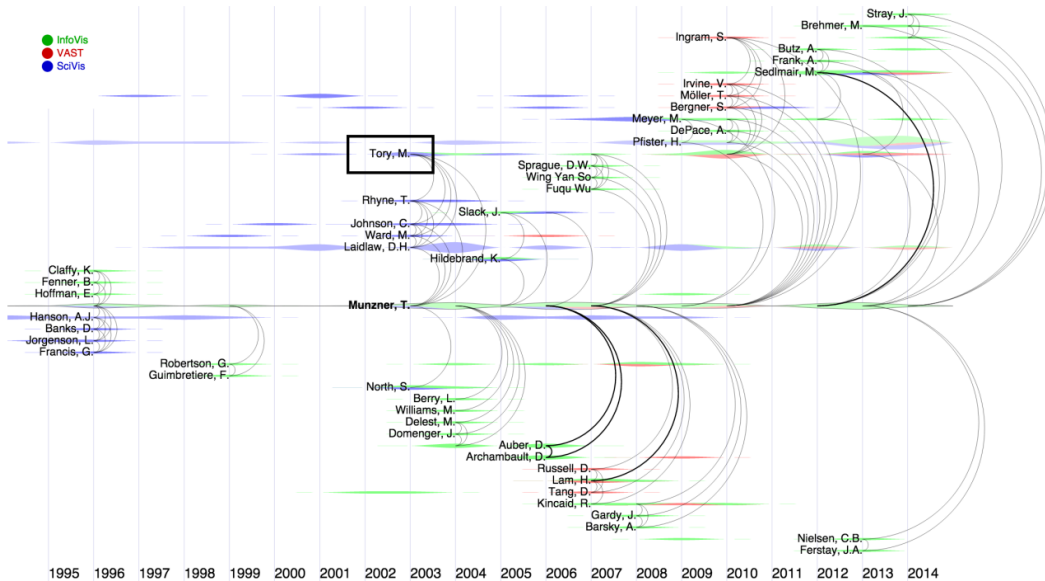


图 3: TimeArcs具体案例

- 16eurovis *A Visual Analytics Framework for Microblog Data Analysis at Multiple Scales of Aggregation*. 利用社交数据/微博数据进行环境感知, 危机感知或者说异常检测等. 文章提到比较厉害的一个技术是context preserving data navigation, 在文中体现的是交互中的一个transparent transition 如图 4, 在过度时这一时间点与下一时间点的隐现上做文章. 做微博分析的很多, 文中还概括了一下微博挖掘的问题; 此外在话题分类时采用了一种算法, 专用于crisis-related microblogs.

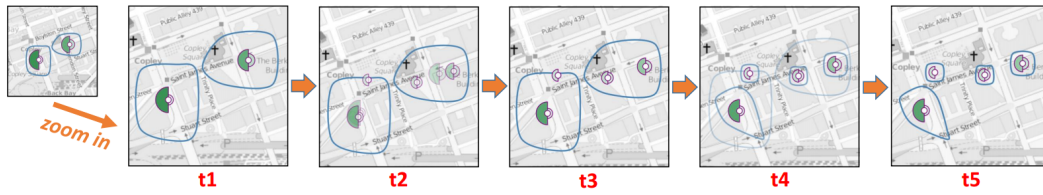


图 4: transparent transition技术